6-1-2005

# A Comparative Evaluation of Spectral Quality Metrics for Hyperspectral Imagery

John P. Kerekes
*Rochester Institute of Technology*

Adam Cisz
*Rochester Institute of Technology*

Rulon Simmons
*ITT Industries*



Recommended Citation
John P. Kerekes, Adam P. Cisz, Rulon E. Simmons, "A comparative evaluation of spectral quality metrics for hyperspectral imagery", Proc. SPIE 5806, Algorithms and Technologies for Multispectral, Hyperspectral, and Ultraspectral Imagery XI, (1 June 2005); doi: 10.1117/12.605916; https://doi.org/10.1117/12.605916

# A comparative evaluation of spectral quality metrics for hyperspectral imagery

John P. Kerekes[*a], Adam P. Cisz[a], Rulon E. Simmons[b]
[a]Center for Imaging Science, Rochester Institute of Technology, Rochester, NY 14623
[b]Space Systems Division, ITT Industries, Rochester, NY

## ABSTRACT

Quantitative methods to assess or predict the quality of a spectral image are the subject of a number of current research activities. An accepted methodology would be highly desirable to use for data collection tasking or data archive searches in way analogous to the current uses of the National Imagery Interpretation Rating Scale (NIIRS) General Image Quality Equation (GIQE). A number of approaches to the estimation of quality of a spectral image have been published. An issue with many of these approaches is that they tend to be constructed around specific tasks (target detection, background classification, etc.) While this has often been necessary to make the quality assessment tractable, it is desirable to have a method that is more general. One such general approach is presented in a companion paper (Simmons, et al[1]). This new approach seeks to get at the heart of the general spectral imagery quality analysis problem – assessing the confidence of an image analyst in performing a specified task with a specific spectral image. In this approach the quality from spatial and spectral aspects of the imagery are treated separately and then a fusion concept known as "semantic transformation" is used to combine the utility, or confidence, from these two aspects into an overall quality metric. This paper compares and contrasts the various methods published in the literature with this new General Spectral Utility Metric (GSUM). In particular, the methods are applied to a target detection problem using data from the airborne HYDICE instrument collected at Forest Radiance I. While the GSUM approach is seen to lead to intuitively pleasing results, its sensitivity to image parameters was not seen to be consistent with previously published approaches. However, this likely resulted more from limitations of the previous approaches than with problems with GSUM. Further studies with additional spectral imaging applications are recommended along with efforts to integrate a performance predication capability into the GSUM framework.

**Keywords**: Spectral imaging, spectral quality

## 1. INTRODUCTION AND OVERVIEW

The ability to quantitatively assess the quality of a multispectral or hyperspectral image is desirable for many reasons including instrument comparisons and trade studies, image archive retrieval and tasking of data collections. However, the notion of the "quality" of a spectral image will depend upon many disparate factors including characteristics of the scene, the sensor, the algorithms applied, and the desired product. While an image with just a few bands but high spatial resolution may have high quality judged by someone looking at spatial information, an image with many bands but moderate spatial resolution may have even higher quality when judged by an analyst looking at spectral information. It is precisely these tradeoffs that one seeks to quantify in the development of a spectral quality measure.

While several research efforts have made progress in helping to identify and quantify these factors, most have been in the context of a given application in order to make the analysis tractable. A more general metric would be desirable. A companion paper by Simmons, et al[1], has proposed such a general metric based on semantic transformations of the spatial and spectral quality treated as separable quantities.

The emergence of this new, more general metric has motivated this comparative evaluation study. In Section 2 we review various published efforts at quantifying spectral quality. We then describe the data set and analysis approach used in our comparative evaluation in Section 3. Section 4 describes our results and Section 5 provides a brief discussion and interpretation of these results. Finally, in Section 6 we summarize and identify future work.

[*] kerekes@.cis.rit.edu



Algorithms and Technologies for Multispectral, Hyperspectral, and Ultraspectral Imagery XI,
edited by Sylvia S. Shen, Paul E. Lewis, Proceedings of SPIE Vol. 5806
(SPIE, Bellingham, WA, 2005) · 0277-786X/05/$15 · doi: 10.1117/12.605916

## 2. SPECTRAL QUALITY

### 2.1 Spectral Quality Discussion

The notion of the "quality" of a multispectral or hyperspectral image deserves some discussion, as there is not a universally accepted definition of the term. In this usage, the dictionary defines quality as "degree or grade of excellence." Thus, a spectral quality measure should contain a monotonically increasing scale that represents the degree of excellence (in some sense) of a spectral image. The use of a numerical scale to describe the quality is a convenient way of ordering the values and comparing disparate images. It is intuitive that the higher the numerical value, the higher the quality.

Several research groups have explored approaches to the quantification of spectral quality. The next several subsections provide a short description of different approaches identified from the literature.

### 2.2 Martin, Vrabel and Leachtenauer[2]

This group defined quality as "…the extent to which an image or data set precisely replicates the scene represented by the image or data set." They also defined utility separately as "…the value of a product to a user in satisfying information requirements." An approach was outlined in their paper to 1) objectively calculate the quality of an image, 2) subjectively determine the utility through analyst assessments, and then 3) ultimately relate these two metrics to obtain an objective utility metric. They defined eight objective quality metrics: ground sample distance, spatial relative edge response, geolocation accuracy, full width at half maximum of the sensor spectral response function, spectral relative edge response, spectral calibration accuracy, noise equivalent delta radiance and radiometric calibration accuracy.

For the utility metric, they recognized the complexity of the analysis process and presented one example measure as the probability of correct material identification, $P_{CI}$. They postulated $P_{CI}$ to be a function of a number of parameters including the accuracy of signature definition, the sensor performance (spatial, spectral, and radiometric), the analysis of system performance, the sample abundance, and a decision criterion. For this initial effort, their plans were to assess utility through objective questions to a group of analysts, each performing an independent analysis of a spectral image. This is analogous to the methodology of the General Image Quality Equation (GIQE)[3] for the National Imagery Interpretation Rating Scale (NIIRS).

The issues and approach outlined in their work set the stage for follow-on efforts. While no publication could be found to indicate if the work was continued to obtain the objective metric postulated, their efforts deserve recognition as one of the first groups to attempt to tackle this problem. Also, the approach they described served as a model for subsequent efforts.

### 2.3 Sweet, Granahan, and Sharp[4]

Another method of assessing image utility is the Spectral Similarity Value (SSV), patented by Sweet, et al. This approach uses a two-part equation to evaluate the similarity between two spectral vectors based on (1) the magnitude difference and (2) the shape difference. The SSV approach is appropriate for use with hyperspectral images. It combines the calculations of the magnitude difference and the shape difference, giving each equal weighting. This approach may not be valid in general. The SSV approach can only compare a single spectrum with another single spectrum and therefore is not well-suited to take into account variability in either the target or backgrounds.

They defined the SSV as being the root-sum-square of the Euclidean distance and one minus the cross-correlation coefficient between two spectral vectors. Mathematically, for two normalized spectral vectors $x$ and $y$ with $N$ bands,

$$SSV(x,y) = \sqrt{d^2 + \hat{r}^2} \tag{1}$$

where,

$$d = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(x_i - y_i)^2} \tag{2}$$

and,

$$\hat{r}^2 \quad = \quad 1 - \left( \frac{\frac{1}{N}\sum_{i=1}^{N}(x_i - \mu_x)(y_i - \mu_y)}{\sigma_x \sigma_y} \right)^2 \tag{3}$$

where $\mu_x$ and $\mu_y$ are the means and $\sigma_x$ and $\sigma_y$ the standard deviations of $x$ and $y$ computed across the $N$ wavelengths.

**2.4 Kerekes and Hsu[5,6]**

This work was closely modeled after the NIIRS GIQE in that a number of analyses were conducted with spectral images (or model analyses) of varying quality, followed by the development of regression equations relating the spectral image parameters to analysis task performance.

In the first paper[5], model-based trade studies and empirical analyses of matched filter target detection in the reflective VNIR/SWIR spectral region were performed. Regression analysis was then applied to these results and the spectral image parameters deemed to be of most importance to detection performance, resulting in an objective equation for a Spectral Quality Rating Scale (SQRS) value. Slightly different equations were obtained between the model and the empirical analysis, but the relative ordering of images with different parameters remained the same. In our current work, we will use the empirically-derived result for comparison.

$$SQRS_{detection} \quad = \quad 9.70 - 3.32\log_{10}[GSD(cm)] + 0.67\log_{10}[SNR] + 0.48\log_{10}[N] \tag{4}$$

Here, *GSD* is the ground sample distance, *SNR* is the traditional signal-to-noise ratio, and $N$ is the number of spectral channels in the spectral region of interest.

The second paper[6] considered a similar approach using empirical analysis but for the application of terrain classification. Minimum distance and spectral angle metrics were used to classify land cover type for a desert and a forest scene. The performance metric selected was the Kappa statistic, which is a scalar representation of the traditional classification confusion matrix. Equations of the following form were developed with a regression between the Kappa statistic and the image parameters.

$$Kappa \quad = \quad A + B\log_{10}[GSD(cm)] + C\log_{10}[SNR] + D\log_{10}[N] \tag{5}$$

The coefficients $A$, $B$, $C$, and $D$ varied quite a bit, depending on the classification algorithm and the scene type. However, they all showed a general trend of higher Kappa value (higher performance) with greater GSD, which was the opposite of the target detection result. While smaller GSD (higher spatial resolution) is usually desirable for target detection (greater numbers of pixels or greater target fraction of pixel in the unresolved case), larger GSD's can lead to improved land cover classification performance due to an averaging of the within-class variation. These opposite trends presented a problem for reconciliation of the two approaches, and suggests the concept that coefficients and possibly even the form of the equation should depend upon the analysis task at hand.

Subsequent to these efforts, Hsu extended the target detection scale to a form particularly well-suited to matched filter detection and taking into account a measure of the separability of the target and the background [personal communication].

$$SQRS_{SCR-detection} \quad = \quad 10.6 - 1.6\log_{10} t - 3.3\log_{10}[GSD(cm)] + 1.6\log_{10}[SCR] \tag{6}$$

Here, *SCR* is the signal-to-clutter ratio defined for a target and background having spectral mean vectors $\mu_t$ and $\mu_b$, and the background having a spectral covariance matrix, $\Sigma_b$, as

$$SCR \quad = \quad \sqrt{(\mu_t - \mu_b)^t \Sigma_b^{-1} (\mu_t - \mu_b)} \tag{7}$$

with $t$ being the threshold on the normalized match filter output test statistic $\theta$ (defined below for a test pixel $x$) that leads to a specified false alarm rate on the image background.

$$\theta(x) = \frac{(\mu_t - \mu_b)^t \Sigma_b^{-1} (x - \mu_b)}{(\mu_t - \mu_b)^t \Sigma_b^{-1} (\mu_t - \mu_b)} \tag{8}$$

The *SCR* formulation accounts for several parameters. Since means and covariance have dimensionality of the number of spectral bands, that parameter is included. Since the background covariance matrix will by nature include the sensor noise, SNR is included. The mean difference and background variability are also included.

**2.5 Shen[8]**

This work also followed the approach of specifying a performance metric, analyzing a large number of images with varying spectral image parameters, and then forming a regression between the metric and parameters. In this case, the metric of interest was detection probability at a specified false alarm rate. For the case of $P_{FA}$ = 0.001, the following equation was found to have a good fit.

$$P_D = 6.25 - 0.81 \log_{10}[GSD(m)] + 0.12 \log_{10}[SNR] - 0.20 \log_{10}[\Delta\lambda(nm)] - 2.43 \log_{10}[\sigma_{scene}] \tag{9}$$

In this equation $\Delta\lambda$ is the average spectral resolution of the channels and $\sigma_{scene}$ is the average across all spectral bands of the standard deviation of the pixels in the scene. The units for $\sigma_{scene}$ were the HYDICE scaled radiance units which are equal to 4/3 μW/cm$^2$-sr-μm. The coefficients were found to vary modestly for other false alarm rates, mostly affecting the constant term as an offset. This formulation has the advantage of considering the scene variability through the $\sigma_{scene}$ parameter.

**2.6 Simmons, et al[1]**

This more general solution to spectral utility prediction is based on two concepts: (1) the transforming of spatial information into a confidence value, and (2) the numerical combining of spatial and spectral confidence values into a single confidence number. The companion paper provides the details while the following paragraphs describe the concept.

In past studies, image analyst confidence values have been related to NIIRS for specific problems called Essential Elements of Information (EEIs). A typical plot of NIIRS versus confidence for a given EEI is a sigmoid-shaped curve, with the curve being different for each EEI. In practice, the relationship between NIIRS and confidence is largely driven by the size of the target relative to the image resolution [11]. In equation form these curves are given as:

$$C_{Spatial} = \frac{(N / N_{50})^E}{1 + (N / N_{50})^E} \tag{10}$$

where:

$E = 2.7 + 0.7(N/N_{50})$,
$N$ = Number of resolution cycles per minimum dimension of target,
$N_{50}$ = Cycle criteria for 50 percent success,

and $N_{50}$ has the following values for detection, recognition, and identification: 1.0±0.25, 4.1±0.35, and 6.4±1.5.

The spectral confidence ($C_{Spectral}$) can be found through an assessment of the separability of target and background spectral distributions or from results of hyperspectral image analysis techniques such as a spectral matched filter. The detection probability at a specified false alarm rate offers one metric that could correspond to confidence. In general, it may be that results from spectral algorithms will require a weighting factor to convert them to a confidence value.

Given the spatial and spectral confidences, it is desirable for them to be combined. This is not simply an additive process, since if you have a confidence of 1.0 (i.e., 100%) from either the spatial or spectral information, you are totally confident of your answer regardless of any further information from the other side. Ideas from development of "semantic transformations" lead to the method shown in (11).

$$C_{Total} = 1-(1-C_{Spatial})\cdot(1-C_{Spectral}) \tag{11}$$

Note that while a spectral image simultaneously gives both spectral and spatial information, this equation is applicable to cases where the information is not coincident in time or space such as fusing information from a high resolution panchromatic image with a lower resolution spectral image taken simultaneously or even at a different time.

## 3. DATA SET AND EVALUATION APPROACH

The data set selected for this comparative evaluation is one that has become a *de facto* standard among hyperspectral processing researchers. It is a set of images collected by the airborne imaging spectrometer HYDICE[9] of a forested region at the U.S. Army Aberdeen Proving Ground in 1995 by the HYMSMO (Hyperspectral MASINT Support to Military Operations) program. Known as Forest Radiance 1 (FR1), the data set contains images collected at varying sensor altitudes (ground resolution) with significant amounts of ground truth (spectra, pictures, etc.) Also, numerous man-made objects were deployed in static positions, offering a wide array of opportunities for target detection analyses.

Figure 1 shows three of the images collected during FR1. They show an open field with man-made objects collected at three different ground resolutions. These data offer an empirical way to study the varying effects due to ground resolution, and to some extent, signal-to-noise ratio.

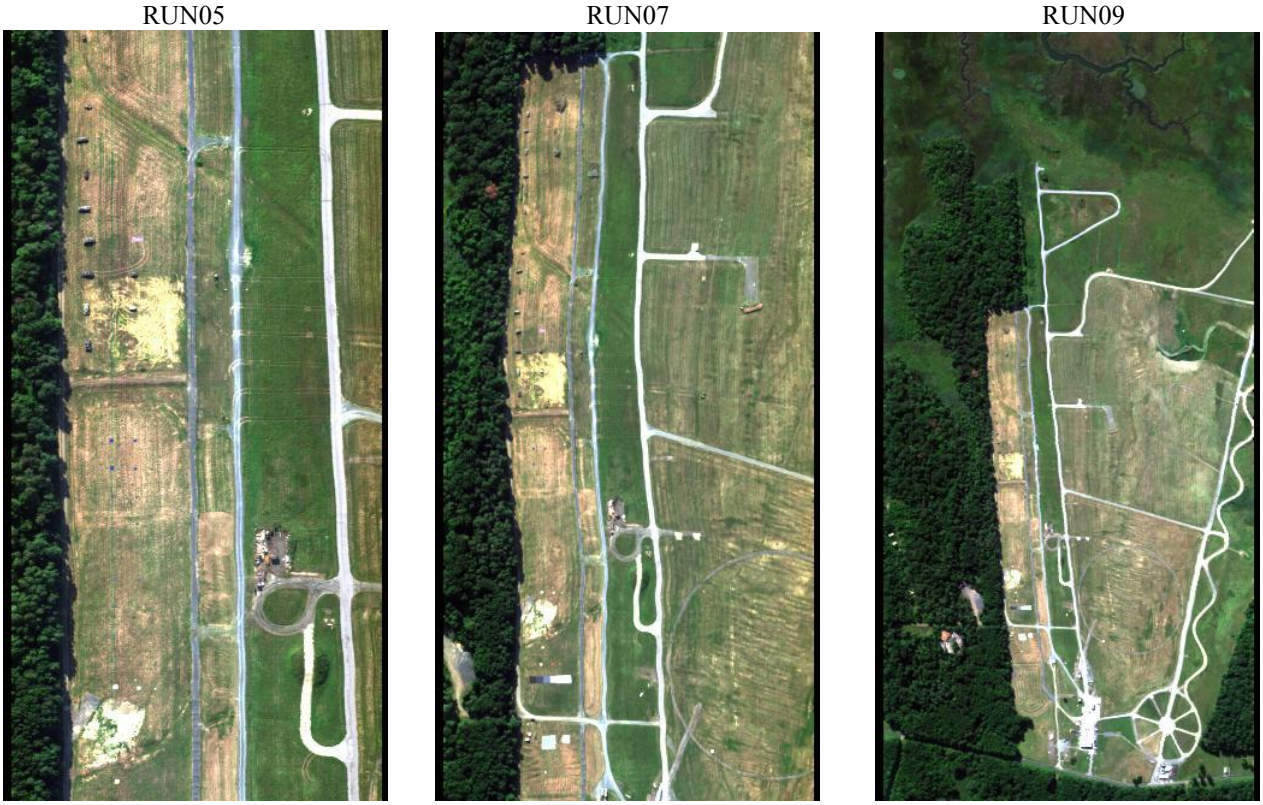


Figure 1. Three images collected during FR1 and used to study effects of sensor parameters on spectral quality.

Table 1 provides the details regarding each of these images. The signal-to-noise ratio (SNR) reported here is the approximate value across atmospheric window spectral regions using a nominal scene radiance and the noise level calculated from calibration sequences as discussed in Nischan, et al[10]. Note for these images, the sensor was found to be fixed noise dominated, and the SNR varies directly with the integration time.

Table 1. Details on FR1 images used in analysis.

| Run | Sensor altitude (m) | GSD (m) | Pixel Integration Time (ms) | SNR |
|---|---|---|---|---|
| 5 | 1570 | 0.75 | 5 | 100 |
| 7 | 3170 | 1.5 | 15 | 300 |
| 9 | 6410 | 3.0 | 15 | 300 |

Images with varying ground resolution, number of spectral channels, and SNR as described in Table 2 were analyzed. In all cases we started with the atmospherically-compensated reflectance data obtained via the Empirical Line Method. Note the 144 channel cases represent the original sensor resolution (eliminating low SNR bands from the 210 bands collected by the instrument). The 36 and 18 channel cases were obtained by aggregating 4 and 8 channels respectively. (The aggregation of 2 channels was omitted since tests showed its results to be very similar to the original resolution.) The lower SNR's (x0.25 and x0.125) were obtained by adding independent Gaussian random noise to the image with a band-varying standard deviation set so the ratio of the mean radiance to the root-sum-square of the additional noise and the original assumed noise equaled the desired SNR. A total of 27 different images were analyzed.

Table 2. Cases of modification to the spectral channels and noise levels for each image. Original in italics.

| Run | GSD (m) | Number of spectral channels | SNR |
|---|---|---|---|
| 5 | 0.75 | *144*, 36, 18 | *100*, 25, 13 |
| 7 | 1.5 | *144*, 36, 18 | *300*, 75, 38 |
| 9 | 3.0 | *144*, 36, 18 | *300*, 75, 38 |

A type of spectral matched filter was applied to the resulting images to obtain detection results for the spectral confidence metric. For selected objects within the scene, we applied the Adaptive Coherence Estimator[12] (ACE) detector. We then estimated the detection and false alarm probabilities on a per-pixel basis to obtain receiver operating characteristic (ROC) curves. At the specified false alarm rate of $10^{-3}$, we obtained the detection probability under each condition and used that as the spectral confidence value, $C_{Spectral}$.

## 4. EVALUATION RESULTS

The various methods of calculating a spectral image quality metric were applied to the 27 HYDICE images to study the sensitivity of the metric value to the image parameters. For each image, we calculated the following metrics.

**SSV**. Calculated using equations (1), (2), and (3) with the mean spectra of the target and the scene-wide background

**SQRS**. Calculated using equation (4) and divided by 10 for comparison purposes.

**SQRS_SCR**. Calculated using equations (6), (7), and (8) with the mean spectrum of the target and the scene-wide background mean spectrum and covariance matrix. The threshold $t$ was selected to achieve $P_{FA} = 10^{-3}$. Also divided by 10 for comparison.

**Pd**. Calculated using equation (9) with a value of 179 (in HYDICE scaled radiance units) for $\sigma_{scene}$ for this forest scene.

**Cspatial**. Calculated using equation (10) with N equal to number of pixels of the minimum dimension of the target..

**Cspectral**. Equal to the detection probability obtained as described above in Section 3 for $P_{FA} = 10^{-3}$.

**Ctotal**. Calculated using equation (11).

Table 3 presents the results for these cases for the images in their original data condition. Four targets in each scene were selected for evaluation. Note that the SQRS and the Pd metrics are independent of the target type and thus are constant within a given image.

Table 3. Spectral quality metrics for the original data condition (no band aggregation nor noise added).

| Scene/Target | SSV | SQRS/10 | SQRS_SCR/10 | Pd | Cspatial | Cspectral | Ctotal |
|---|---|---|---|---|---|---|---|
| Run05/C5 | 0.58 | 0.58 | 0.89 | 0.92 | 0.99 | 0.97 | 1.00 |
| Run05/C6 | 0.43 | 0.58 | 0.80 | 0.92 | 0.99 | 0.92 | 1.00 |
| Run05/V | 0.71 | 0.58 | 0.83 | 0.92 | 0.67 | 0.69 | 0.90 |
| Run05/VF | 0.53 | 0.58 | 0.64 | 0.92 | 0.67 | 0.71 | 0.90 |
| | | | | | | | |
| Run07/C5 | 0.51 | 0.51 | 0.79 | 0.73 | 0.81 | 0.99 | 1.00 |
| Run07/C6 | 0.39 | 0.51 | 0.73 | 0.73 | 0.89 | 0.92 | 0.99 |
| Run07/V | 0.57 | 0.51 | 0.76 | 0.73 | 0.27 | 0.91 | 0.94 |
| Run07VF | 0.46 | 0.51 | 0.61 | 0.73 | 0.10 | 0.78 | 0.80 |
| | | | | | | | |
| Run09/C5 | 0.55 | 0.41 | 0.76 | 0.49 | 0.27 | 1.00 | 1.00 |
| Run09/C6 | 0.39 | 0.41 | 0.67 | 0.49 | 0.27 | 0.95 | 0.97 |
| Run09/V | 0.55 | 0.41 | 0.70 | 0.49 | 0.02 | 1.00 | 1.00 |
| Run09/VF | 0.26 | 0.41 | 0.52 | 0.49 | 0.02 | 0.03 | 0.05 |

Figure 2 plots the various metrics for the V target under the original data condition (all 144 bands and highest SNR). The results in Figure 2 show interesting trends among the various metrics. Since the GSD increases from 0.75 to 1.5 to 3.0m for Run 05, Run 07, and Run 09, respectively, we see that the *Cspatial* metric decreases as one would expect. However, the *Cspectral* metric (and as a result, the *Ctotal* metric) increases. This trend is most likely due to the decrease in variability due to scene averaging as the spatial resolution degrades. This particular target was mostly full pixel. All other metrics also show a downward trend since their values are dominated by GSD (except the SSV). The decrease of the SSV metric is surprising since it is based entirely on the single comparison of the average target and background vectors. It is not clear what caused the downward trend with decreasing spatial resolution for the SSV metric.

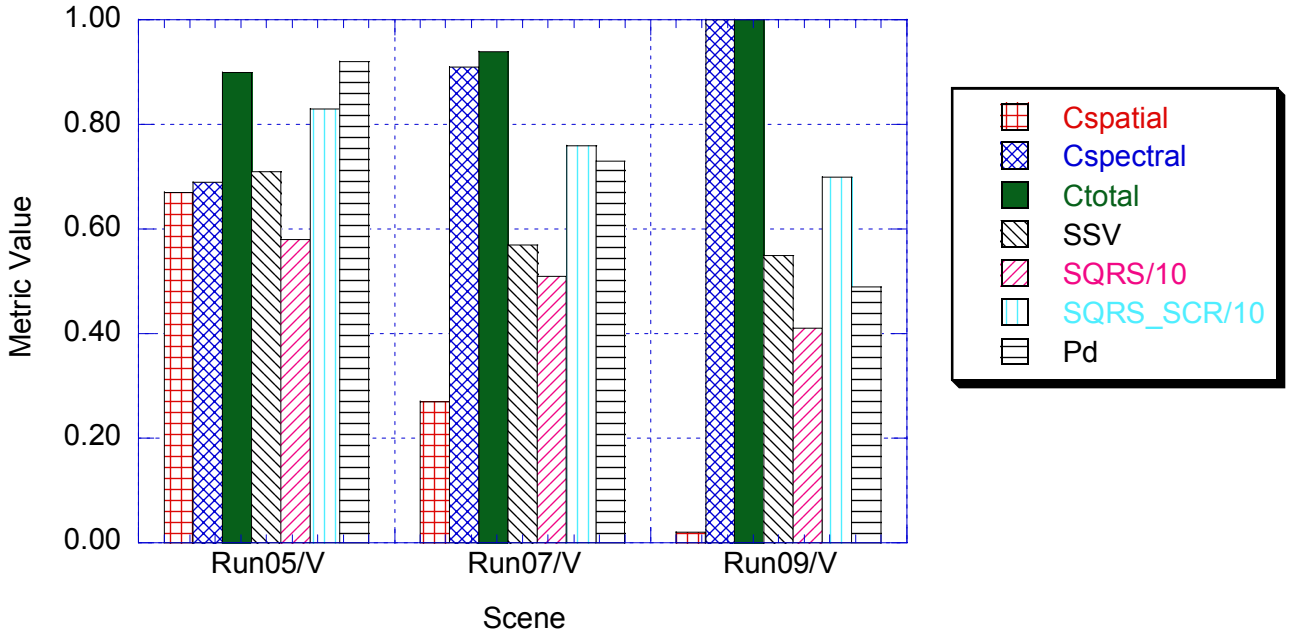


Figure 2. Bar graph of metric results for the V target for the baseline configuration.

Figure 3 presents results showing the trend with SNR for the various spectral metrics. This plot is for the V target and the Run07 GSD (with original number of bands - 144). The *Cspatial* metric is independent of SNR, so it does not change. The SSV metric is calculated using average values, so it also is insensitive to the SNR. The others show the expected slight downward trend, although the *Cspectral* (and as a result, the *Ctotal*) metrics show insignificant changes between 0.25 and 0.125 factors. Thus, the trend is consistent across the metrics in the expected manner, but in general it is rather a small effect even for factors of eight changes in the SNR.

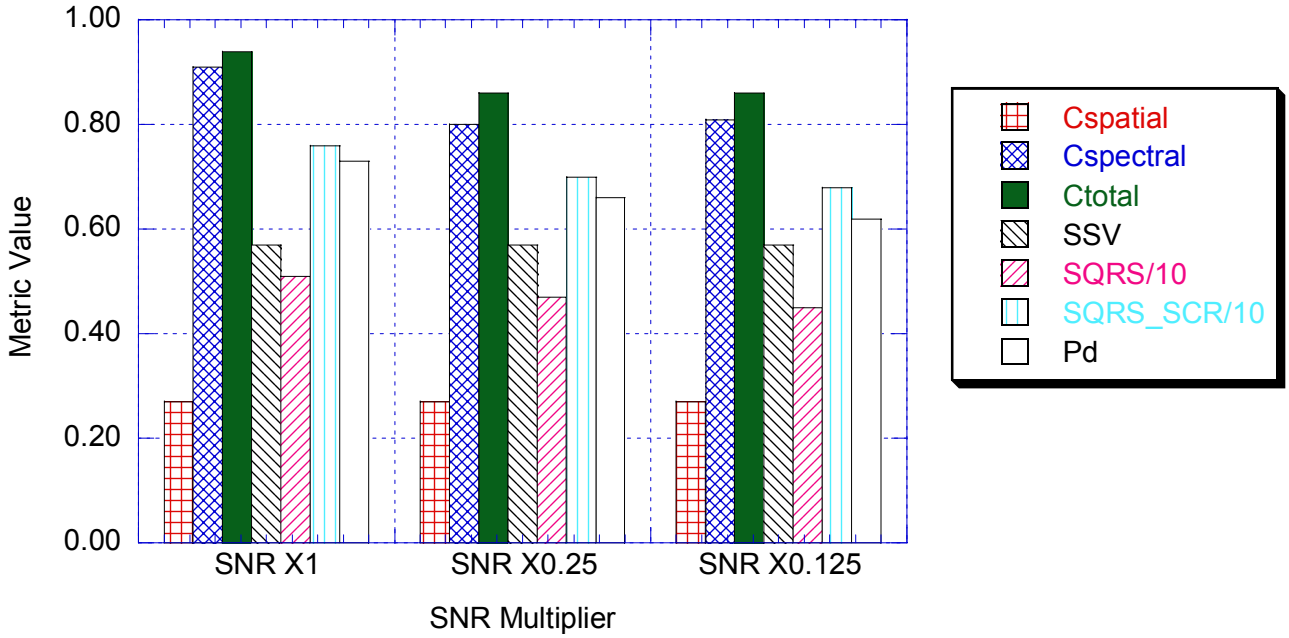


Figure 3. Bar graph of metric results for the V target for various SNR values using Run07.

Figure 4 presents results showing the trend with number of spectral channels for the various spectral metrics. This was done for the V target and the Run07 GSD (with original SNR - 300). The *Cspatial* metric is independent of spectral channels, so it does not change. The SSV metric shows a slight upward trend with fewer channels, while the others show the expected slight downward trend. The contrary result with the SSV metric is again a bit unclear, but could be due to the reduced spectral variability across the channels with the fewer number of channels. Overall, the trend with changes in the number of spectral channels is consistent across the other metrics in the expected manner, and is a bit more noticeable than the trend observed in SNR.

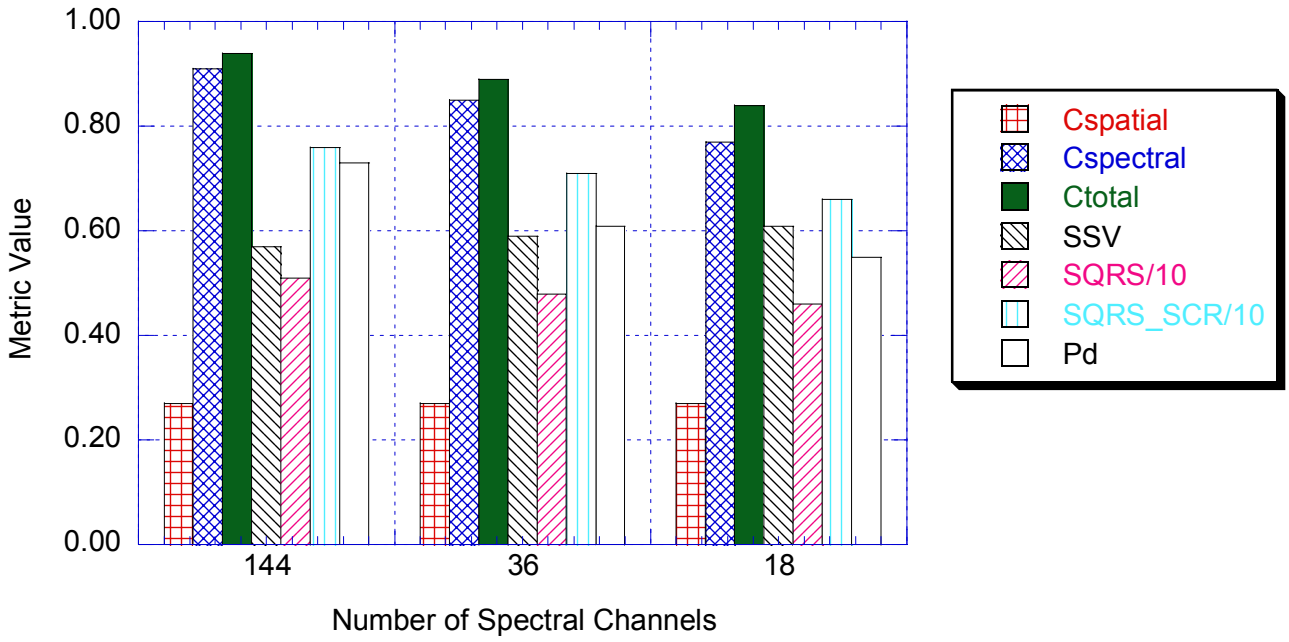


Figure 4. Bar graph of metric results for the V target for various number of spectral channels using Run07.

While these plots show trends that are reasonable for the various metrics, they offer little help in understanding how well they help characterize spectral image quality or predict performance in the analysis of the images. One may reasonably ask: which metrics are most consistent in the performance characterization? To explore this question, we consider results for all targets, all configurations, and all metrics. However, we would like to select one metric for comparison to the others to explore their consistency.

So, how do we select which metric to use as the "standard"? In this study, the spectral confidence metric, *Cspectral*, was the resulting detection probability for a matched filter detection algorithm applied to the data. Thus, it represents a "true" measure of detection performance, at least for the empirical results. Following the methodology of Simmons, et al[1], this was combined with information about the spatial dimension of the problem at hand to form the *Ctotal* metric. Since *Ctotal* was obtained from the empirical analysis modified by descriptive spatial information, it was selected as the standard by which to compare the others.

Figure 5 presents scatter plots between *Ctotal* and the other metrics. Since *Ctotal* can saturate at 1.0, the cases where *Ctotal* = 1.0 were eliminated. This eliminated 41 of the 108 cases (4 targets * 27 images) leaving the 67 cases plotted. A linear curve was fit to each plot and the regression coefficient is given above each plot.

None of these metrics show a high correlation to the selected standard. The SQRS_SCR metric, which included information about the target/background contrast in addition to terms representing the spatial resolution (and implicitly the number of channels and SNR), had the highest regression coefficient. This implies it is the most consistent with *Ctotal*. But, even as the highest, it did not show a tight coupling between performance predicted by its equation and that achieved as captured by the *Ctotal* metric.

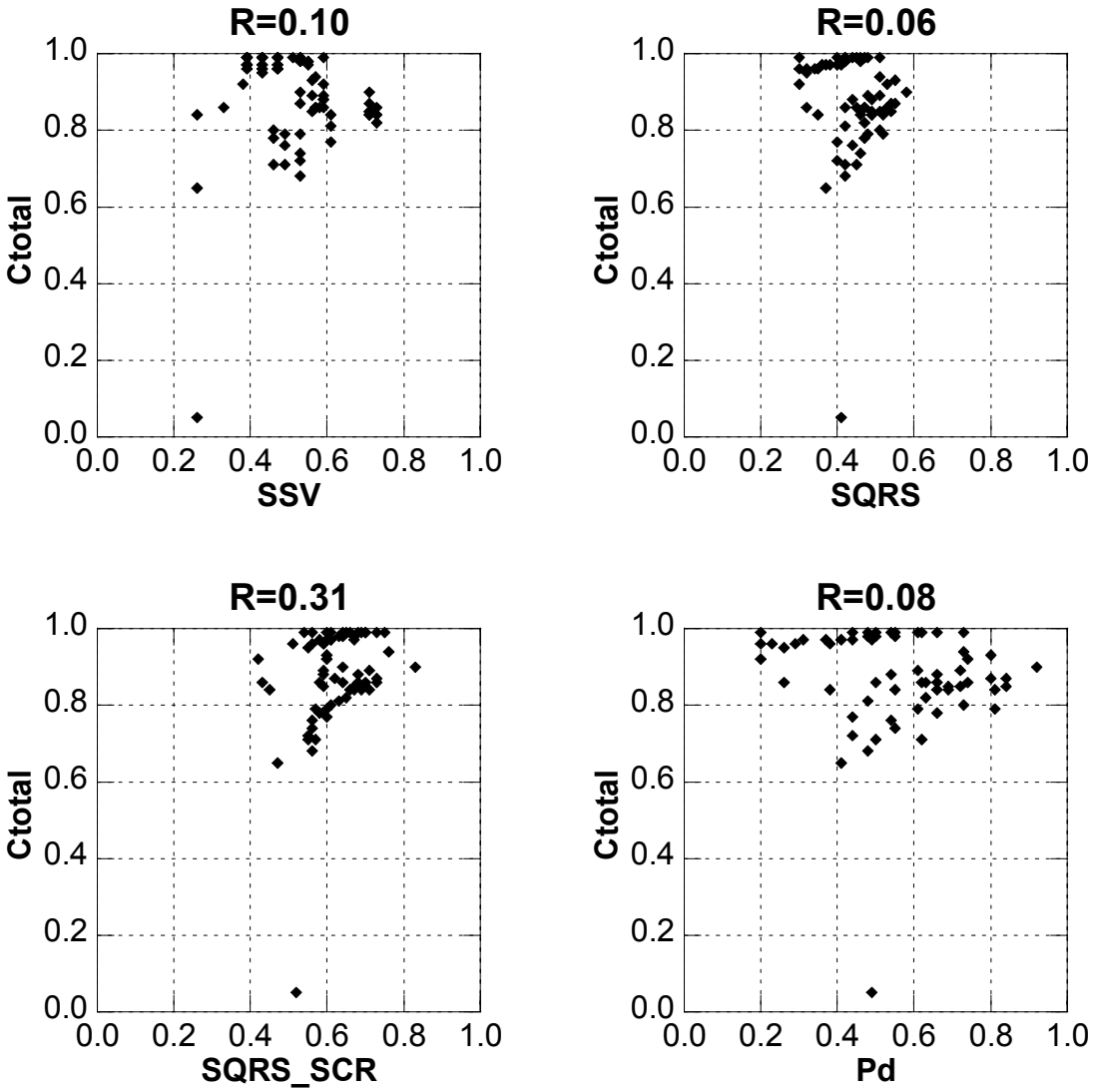


Figure 5. Scatter plots comparing the various spectral metrics to the total confidence metric.

## 5. DISCUSSION

Overall, these results show trends that one expects to see. That is, the spectral image quality metrics tend to increase with higher spatial and spectral resolution, and with higher SNR. However, the scatter between the metrics is quite large demonstrating that the sensitivity to image (sensor) parameters varies quite a bit from metric to metric.

The *Ctotal* metric was selected as the most representative of detection performance since it was based on empirically-obtained detection results and geometric information about the image and the target. Comparison with other previously metrics showed none that were significantly consistent with its sensitivity to spectral image parameters. However, since the spectral confidence component is based on empirically-derived results, its use as a predictive tool is limited.

Table 4 provides a summary comparative evaluation of the various metrics studied. In the top part it identifies which metrics have each of the desirable characteristics of a spectral quality metric. As can be seen, the GSUM *Ctotal* confidence metric is the only one listed that includes all five. However, the predictive capability is derived only from

the spatial side, not the spectral, and that is a current limitation. But its general framework does support the eventual inclusion of such a capability.

Table 4. Comparative summary for spectral quality metric characteristics and parameter sensitivity.

| | SSV | SQRS | SQRS_SCR | Pd | Cspatial | Cspectral | Ctotal |
|---|---|---|---|---|---|---|---|
| **Characteristics:** | | | | | | | |
| Predictive | | X | | X | X | | X |
| Function of image data | X | | X | | | X | X |
| Function of sensor parameters | | X | X | X | X | | X |
| Task dependent | | X | X | X | | X | X |
| Incorporates analyst results | | | | | | X | X |
| | | | | | | | |
| **Parameter Sensitivity:** | | | | | | | |
| Smaller GSD | ↑ | ↑ | ↑ | ↑ | ↑ | ↓ | ↓ |
| Higher SNR | - | ↑ | ↑ | ↑ | - | ↑ | ↑ |
| More Channels | ↓ | ↑ | ↑ | ↑ | - | ↑ | ↑ |

The lower part of Table 4 includes a summary of the sensitivity trends observed for the three dominant parameters considered. For each parameter trend, the resulting metric value trend is shown as an upward arrow if it increases, a downward arrow if it decreases, and a dash if it is insensitive to the parameter. As one can see, there is general agreement among the various metrics. One difference that stands out is the opposite trend for the *Cspectral* (and *Ctotal*) metric for the sensitivity to GSD. This was attributed to the increasing variability (and resulting increased target/background distribution overlap) resulting from the increased heterogeneity with smaller GSD's. Since the targets were mostly full pixel, this effect dominated in the empirical results rather than the effects due to increased target fill factor for sub-pixel targets.

What does all this mean? Perhaps the best interpretation is that there are still opportunities out there for researchers to develop robust spectral image quality metrics.

## 6. SUMMARY AND FUTURE WORK

A comparative evaluation has been initiated to examine differences among published spectral image quality metrics. Sample HYDICE imagery were used to empirically evaluate the various metrics. Most previous metrics were calculated in the context of a single task, without much regard to all the parameters. The metric proposed by Simmons, et al, the General Spectral Utility Metric (GSUM), *Ctotal*, takes into account both spectral and spatial information, in the context of the task at hand. While it seems intuitively the most comprehensive of the metrics evaluated, it remains to be seen whether it can evolve into the robust metric desired by the community.

As a next step, the GSUM approach should be evaluated for other spectral imagery analysis tasks including land classification, and material identification and quantification. Ultimately, it will need to be coupled to some form of performance prediction model to enable its use as a predictive tool to support collection planning and spectral image archival search applications.

## REFERENCES


1. R. Simmons, T. Elder, D. Stewart, E. Cincotta, C. Kennedy, and R. Craig Van Nostrand, "General Spectral Utility Metric for Spectral Imagery," *Proceedings of Algorithms and Technologies for Multispectral, Hyperspectral, and Ultraspectral Imagery XI*, SPIE Vol. 5806, 2005.
2. L. Martin, J. Vrabel, J. Leachtenauer, "Metrics for Assessment of Hyperspectral Image Quality and Utility," *Proceedings of International Symposium on Spectral Sensing Research*, 1999.
3. J. Leachtenauer, W. Malila, J. Irvine, L. Colburn, and N. Salvaggio, "The General Image Quality Equation," *Applied Optics*, pp. 8322-8328, 10 November 1997.
4. J. Sweet, J. Granahan, and M. Sharp, "An Objective Standard for Hyperspectral Image Quality," *Proceedings of AVIRIS Workshop*, Jet Propulsion Laboratory, Pasadena, California, 2000.
5. J. Kerekes and S. Hsu, "Spectral Quality Metrics for VNIR and SWIR Hyperspectral Imagery," *Proceedings of Algorithms and Technologies for Multispectral, Hyperspectral, and Ultraspectral Imagery X*, SPIE Vol. 5425, 2004.
6. J. Kerekes and S. Hsu, "Spectral Quality Metrics for Terrain Classification," *Proceedings of Imaging Spectrometry X*, SPIE Vol. 5546, 2004.
7. J. Kerekes and J. Baum, "Spectral Imaging System Analytical Model for Subpixel Object Detection," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 40, no. 5, pp. 1088-1101, May 2002.
8. S. Shen, "Spectral Quality Equation Relating Collection Parameters to Object/Anomaly Detection Performance," *Proceedings of Algorithms and Technologies for Multispectral, Hyperspectral, and Ultraspectral Imagery IX*, SPIE Vol. 5093, 2003.
9. L. Rickard, R. Basedow, E. Zalewski, P. Silverglate, and M. Landers, "HYDICE: An Airborne System for Hyperspectral Imaging," *Proceedings of Imaging Spectrometry of the Terrestrial Environment*, SPIE Vol. 1937, pp. 173-179, 1993.
10. M. L. Nischan, J. P. Kerekes, J. E. Baum, R. W. Basedow, "Analysis of HYDICE Noise Characteristics and Their Impact on Subpixel Object Detection," *Proceedings of Imaging Spectrometry V*, SPIE Vol. 3753, 1999.
11. L. Biberman, Editor, "Electro-Optical Imaging: System Performance and Modeling," SPIE Press Monograph Vol. PM96, SPIE - International Society for Optical Engineering, 2001.
12. S. Kraut and L. Scharf,, "The CFAR adaptive sub-space detector is a scale-invariant GLRT," *IEEE Transactions on Signal Processing*, vol. 47, pp. 2538-2541, September 1999.